MANUEL D'INSTALLATION de GEONETWORK sur un serveur LINUX CentOS 4

	Nom	Date
Etablie par	**Guillain Assoumou** **Equipe IS**	31/08/2007
Vérifiée par		
Validée par		

Avertissement :

Ce document a été **réalisé par un étudiant** de 2ème année de l'EPSI de Montpellier, Guillain Assoumou, dans le cadre d'un stage au sein de l'équipe 'informatique scientifique' DSI* de l'IRD* sous la responsabilité de Régis Hocdé & Bruno Granouillac (de juin à août 2006).

Il s'agit de la **restitution de l'étudiant 'non corrigée'**.

SOMMAIRE

Sommaire

Introduction 3

I°) JAVA 5

II°) APACHE 8

III°) TOMCAT 8

IV°) Installation GEONETWORK / GEOSOURCE 13

V°) Config Harvesting 19

Introduction

Cette documentation indique la manière d'installer Geonetwork version 2.0.3 sur un serveur linux centOS 4 avec comme autres composants annexes :

- le servlet TOMCAT 6.XX
- le serveur apache
- le serveur BDD Mysql

Il est utile de savoir que l'installation dont traite cette documentation a été réalisé sur les serveurs de l'IRD de Montpellier et en tout premier lieu sur le serveur **http:// vmmdweb.pdmz.mpl.ird.fr.**

Configuration matérielle minimale

La configuration minimale requise pour votre ordinateur afin d'utiliser au mieux les performances de Géosource est la suivante :

- Un PC équipé d'un processeur cadencé à 300 MHz minimum (recommandé) ou à
- 233 MHz (requis) (système mono ou bi-processeur), famille Pentium/Celeron, famille
- AMD K6/Athlon/Duron ou compatible recommandé
- 1 Go de mémoire vive au minimum recommandé (512 Mo minimum requis; risque de nuire aux performances et à certaines fonctionnalités)
- 70 Mo d'espace libre au minimum pour l'installation de Géosource

PREREQUIS

- Un navigateur Internet compatible FireFox 1.5+ ou Internet Explorer 6+. Aucun plugin n'est nécessaire.
- Un serveur de base de données accessible avec des droits de création de tables et d'ajout / modification / suppression des données.
- Un serveur Web.
- Un serveur d'application de servlets.
- Un serveur de mail SMTP.

La configuration de leur poste personnel pour pourvoir configurer et exploiter le serveur

Un serveur virtuel basé sur un serveur linux avec des accès forcément distants

Un dépôt WEBDAV qui permettait de déposer des fichiers sur le serveur

Une interface de système de gestion de fichiers Webmin qui permettait également de réaliser quelques commandes système.

Un accès SSH a par la suite été accordé pour faciliter l'accès au serveur.

L'accès SSH a été exploitée via l'outil CYGWIN qui permet d'émuler un environnement linux sur une distribution Windows.

Configuration de CYGWIN

Après avoir installé Cygwin, il est nécessaire de le configurer en téléchargeant les packages nécessaires à son fonctionnement.

La première fois et à chaque fois que vous voudrez paramétrer le logiciel CYGWIN, vous devrez lancer le fichier.exe responsable de son installation.

Il vous est nécessaire de télécharger et d'installer les packages X11 (toute la catégorie) , mais également les packages openssh et inetutils .

Au préalable, il faut instancier la variable DISPLAY

DISPLAY 127.0.0.1 :0.0

Puis de l'exporter

Export DISPLAY

Une fois installé , l'accès au serveur via CYGWIN se fait en lançant la commande :

ssh login@nomServeur ➔ mot de passe.

Une fois logué sur le serveur, connectez vous en root

I°) JAVA

Le but final est d'installer Geonetwork or cette dernière application est développé en JAVA et elle nécessite un environnement JAVA afin de fonctionner.

Les environnements JAVA sont disponibles sur http://java.sun.com/

Mode opératoire manuel

Il est nécessaire de prendre une version java récente, dans tous les cas, supérieure à **Java SDK version 1.5.0.7.**

Pour connaître la version de l'environnement d'exécution JAVA, entrez la commande :
Java –version

Télécharger un RPM adéquat qui s'extrait automatique (« self-extracting »)

Exécuter la commande **rpm –i java_development_kit ou java_runtime_environment** dans le dossier **/user/java.**

*Dans le dossier **/etc/profile.d/***

Vous pouvez faire un fichier java.sh comportant les lignes :

Export JAVA_HOME=/usr/lib/java/latest

Export CATALINA_HOME = tomcat

Mode opératoire automatique

Avec yum :

Faire :

yum list *jdk*

Pour connaitre les versions disponibles sur les dépôts visités par la commande yum de votre environnement

une fois une version repérée dans la liste qui vous est apparue . faire un

yum intall du package désiré.

Mode opératoire automatique destinée aux personnes travaillant dans l'intranet de l'IRD Montpellier

Le jdk de SUN

wget http://vmsiladmin.intranet.mpl.ird.fr/jdk-6u2-linux-i586-rpm.bin

chmod +x jdk-6u2-linux-rpm.bin

./jdk-6u2-linux-rpm.bin

ATTENTION aux problèmes

Source d'erreurs possible, si plusieurs versions de java sont présentes sur votre système, il risque d'y avoir des conflits entre ces dernières. C'est pour cela qu'il est nécessaire, si besoin d'installation de package java il y a, d'enlever les versions obsolètes avant d'installer la nouvelle. (yum remove)

Une fois les variables d'environnement JAVA affectées, vous pouvez vérifier si tout s'est bien déroulé en tapant la commande **env** qui répertorie les variables d'environnement assignées

Les utilisateurs, munis de la version 1.6 de l'environnement d'exécution JAVA (Java Runtime Execution), DOIVENT télécharger le fichier parser_xalan.zip sur le projet Geonetwork du site http://admisource.gouv.fr/ et le décompresser dans le répertoire <install JRE1.6>\lib\ext\,.
Cette archive contient deux fichiers jar :
Xalan.jar version 2.7.0
Serializer.jar version 2.7.0

II°) APACHE

La présence d'un serveur apache est nécessaire pour la publication des pages web car il s'agit d'un serveur web.

Les binaires pour Windows et les sources pour Unix sont disponibles ici : http://httpd.apache.org/download.cgi.

Sous Unix/Linux il faudra soit le recompiler, soit trouver un package ou un rpm tout prêt.

C'est le SIL qui a la responsabilité d'installer Apache sur le serveur sur lequel vous souhaitez travailler.

III°) TOMCAT

Qu'est ce que tomcat ?

Tomcat est un servlet java pour interpréter les pages jsp.

Apache Tomcat est un conteneur de servlet J2EE. Issu du projet Jakarta. Tomcat implémente les spécifications des servlets et des JSP de Sun Microsystems. Il inclut des outils pour la configuration et la gestion, mais peut également être configuré en éditant des fichiers de configuration XML. Comme Tomcat inclut un serveur HTTP interne, il est aussi considéré comme un serveur HTTP.

Pour le faire automatiquement :

um list *tomcat*

yum instamm version_tomcat (vous pouvez prendre la dernière version en ligne)

sinon pour une installation manuel

aller sur le site http://tomcat.apache.org/.

Télécharger la dernière version de Tomcat (supérieur à la 4.0)

Mode opératoire automatique destinée aux personnes travaillant dans l'intranet de l'IRD Montpellier

Installation des RPMs

TOMCAT

Le package tomcat créé dans le cadre de l'offre IS:

rpm -Uvh http://vmsiladmin.intranet.mpl.ird.fr/tomcat-6.0.14-1.mpl.i386.rpm

On configure tomcat pour se lancer automatiquement au démarrage du serveur:

chkconfig –add tomcat

MOD_JK

Le package mod_jk créé dans le cadre de l'offre IS:

rpm -Uvh http://vmsiladmin.intranet.mpl.ird.fr/mod_jk-1.2.25-1.mpl.i386.rpm

Configuration

Mod_jk:

- On créé les fichiers de logs du mod_jk:

  # touch /var/log/httpd/mod_jk.log

  # touch /var/log/httpd/shm.log

- Le module mod_jk se configure à l'aide de deux fichiers:

  - **/etc/httpd/conf.d/mod_jk.conf**

On édite le fichier **/etc/httpd/conf.d/mod_jk.conf** pour y ajouter les lignes suivantes:

```
JkWorkersFile /etc/httpd/conf/workers.properties
JkLogFile /var/log/httpd/mod_jk.log
JkLogLevel error
JkMount /*.jsp ajp13
JkShmFile /var/log/httpd/shm.log
```

```
JkLogLevel info

JkLogStampFormat "[%a %b %d %H:%M:%S %Y] "

# Send everything for context /examples to worker named worker1 (ajp13)

JkMount /examples/* worker1
```

❖ /etc/httpd/conf/workers.properties

On copie le fichier workers.properties par défaut:

 # cp -a /etc/httpd/conf/workers.properties /etc/httpd/conf/workers.properties.DIST

On édite un fichier **/etc/httpd/conf/workers.properties** avec les lignes suivantes:

```
worker.list=worker1

# Set properties for worker1 (ajp13)

worker.worker1.type=ajp13

worker.worker1.host=localhost

worker.worker1.port=8009
```

On redémarre le service httpd pour que les modifications soient prises en compte:

 # service httpd restart

Annexes

mod_jk.conf:

```
LoadModule jk_module modules/mod_jk.so

JkWorkersFile /etc/httpd/conf/workers.properties

JkLogFile /var/log/httpd/mod_jk.log

JkLogLevel error

JkShmFile /var/log/httpd/shm.log

JkLogStampFormat "[%a %b %d %H:%M:%S %Y] "

# Send everything for context /examples to worker named worker1 (ajp13)

JkMount /examples/* worker1
```

workers.properties:

```
worker.list=worker1

# Set properties for worker1 (ajp13)

worker.worker1.type=ajp13

worker.worker1.host=localhost

worker.worker1.port=8009
```

IV°) Installation GEONETWORK / GEOSOURCE

Installation en version serveur

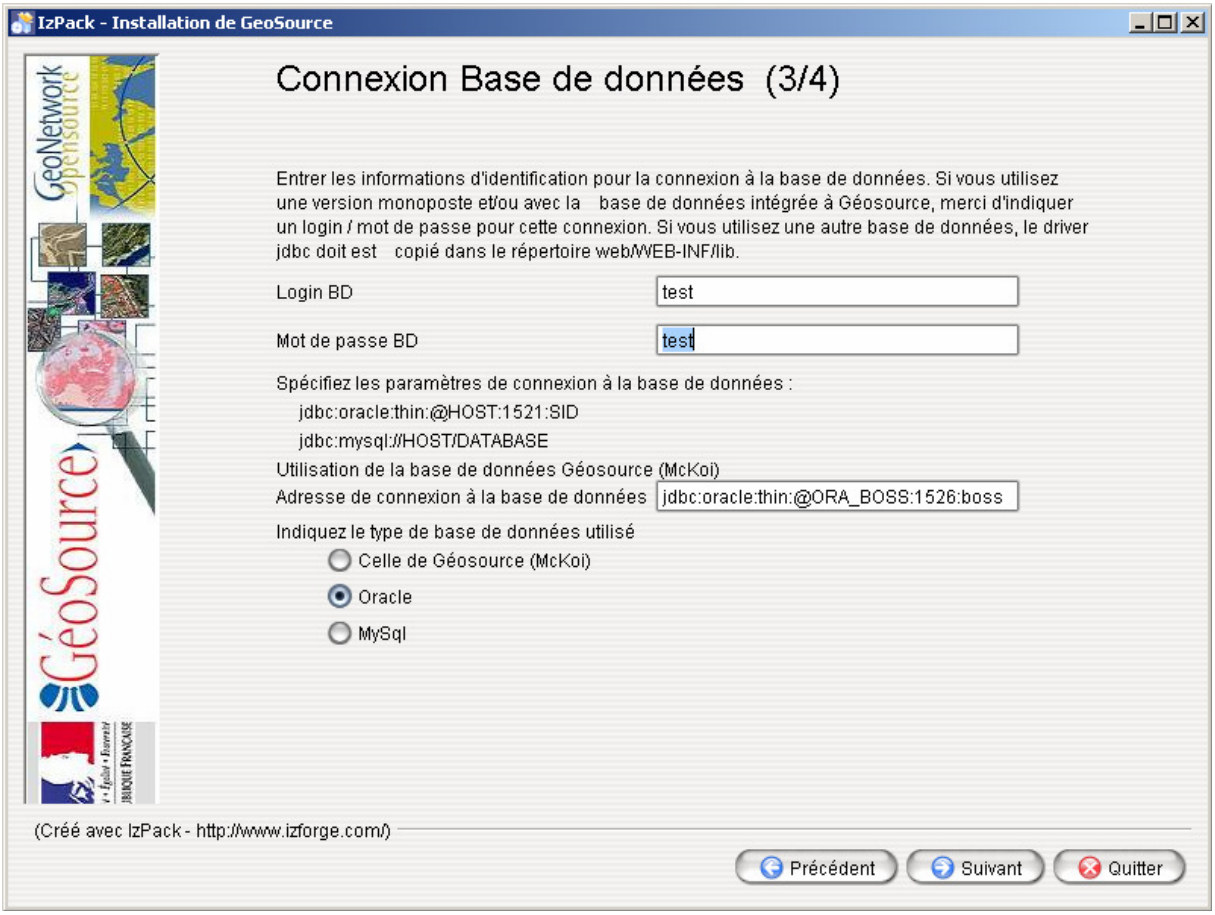
PRINCIPES

On appelle l'installation « serveur » toute installation n'utilisant pas l'ensemble des composants de base de l'application, i.e. jetty, McKoi.
Toutes les combinaisons sont possibles en terme d'installation : Jetty / BD Oracle, Tomcat / McKoi,…

LE SERVEUR DE BASE DE DONNEES

Une base de données relationnelle est indispensable pour Géosource. Actuellement, il existe trois « scripts » disponibles pour l'installation :
- pour McKoi : [répertoire d'installation]\setup\sql\create-db-mckoi.sql
- pour Oracle : [répertoire d'installation]\setup\sql\create-db-oracle.sql
- pour MySQL : [répertoire d'installation]\setup\sql\create-db-mysql.sql

Néanmoins, ces scripts peuvent être adaptés pour d'autres bases de données.
La procédure est la suivante avec l'installeur Géosource :
- Installer un serveur de base de données ;
- Paramétrer le serveur de base pour autoriser les accès JDBC sur un port donné [par exemple, localhost:1521 pour Oracle sur la même machine que Géosource]
- Créer un utilisateur ayant des droits de création de tables ;
- Dans l'écran de choix de base de données, indiquez le login / password de cet utilisateur et paramétrez la configuration avec le driver de VOTRE base de données en précisant l'adresse de localisation de la base de données et son port


IzPack - Installation de GeoSource
Connexion Base de données (3/4)
Entrer les informations d'identification pour la connexion à la base de données. Si vous utilisez une version monoposte et/ou avec la base de données intégrée à Géosource, merci d'indiquer un login / mot de passe pour cette connexion. Si vous utilisez une autre base de données, le driver jdbc doit est copié dans le répertoire web/WEB-INF/lib.
Login BD
test
Mot de passe BD
test
Spécifiez les paramètres de connexion à la base de données :
jdbc:oracle:thin:@HOST:1521:SID
jdbc:mysql://HOST/DATABASE
Utilisation de la base de données Géosource (McKoi)
Adresse de connexion à la base de données
jdbc:oracle:thin:@ORA_BOSS:1526:boss
Indiquez le type de base de données utilisé
Celle de Géosource (McKoi)
Oracle
MySql
(Créé avec IzPack - http://www.izforge.com/)
Précédent
Suivant
Quitter

- Finaliser l'installation.

Cette procédure crée le schéma de la base de données selon les scripts précédents et elle « remplit » la base des données initiale.

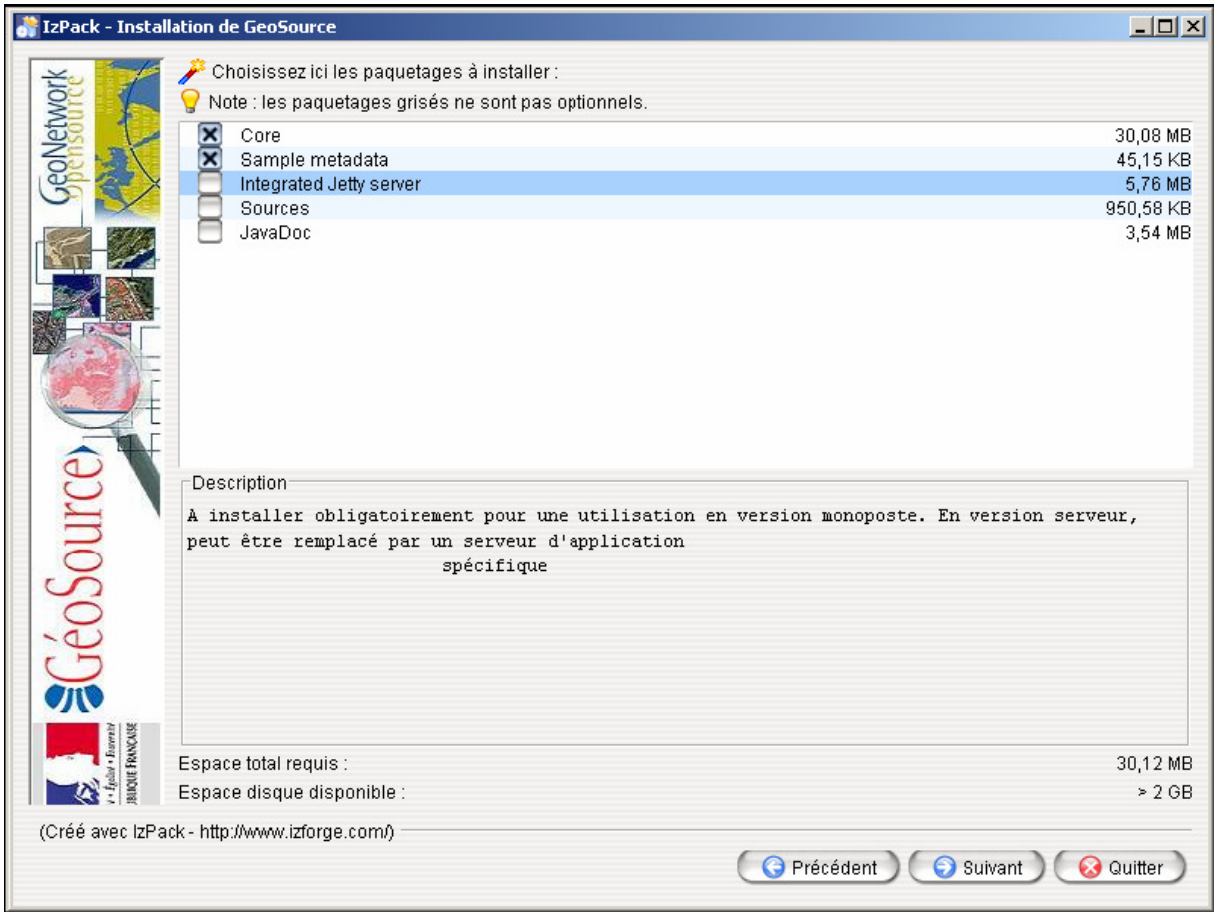
Ces données sont disponibles dans [Répertoire d'installation]\setup\db au format ddf.

Pour des raisons de sécurité de votre base de données, vous pouvez modifier dans Géosource l'utilisateur accédant aux tables. En effet, passé l'installation, l'utilisateur de la base de données a uniquement besoin de droits de lecture / écriture sur les tables. Ouvrez le fichier WEBINF/ config.xml et recherchez la balise <resource enabled="true"> (attention, les balises à false ne sont pas actives). Modifiez le contenu des balises <user>[LOGIN]</user> et <password>[Password]</password>.

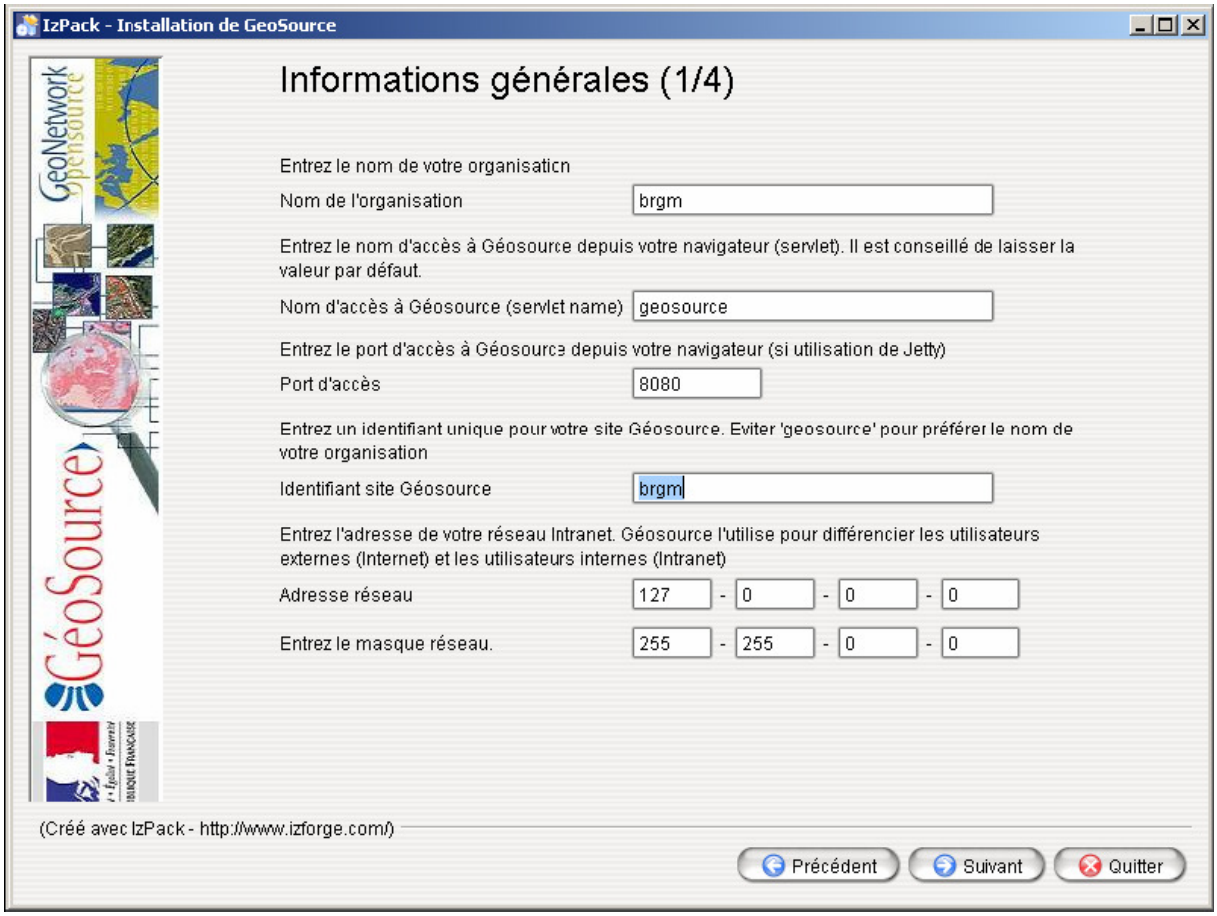
LE SERVEUR WEB ET LE SERVEUR D'APPLICATION

Un serveur d'application et un serveur web sont nécessaires pour Géosource. Les tests ont été réalisés sur Jetty et Tomcat.
Le serveur Tomcat peut être utilisé pour Géosource selon l'une des architectures suivantes :
- Tomcat avec son propre serveur Web, généralement accessible via le port 8080 ;
- Tomcat connecté avec le serveur Web Apache,
- Tomcat connecté avec le serveur Web IIS.

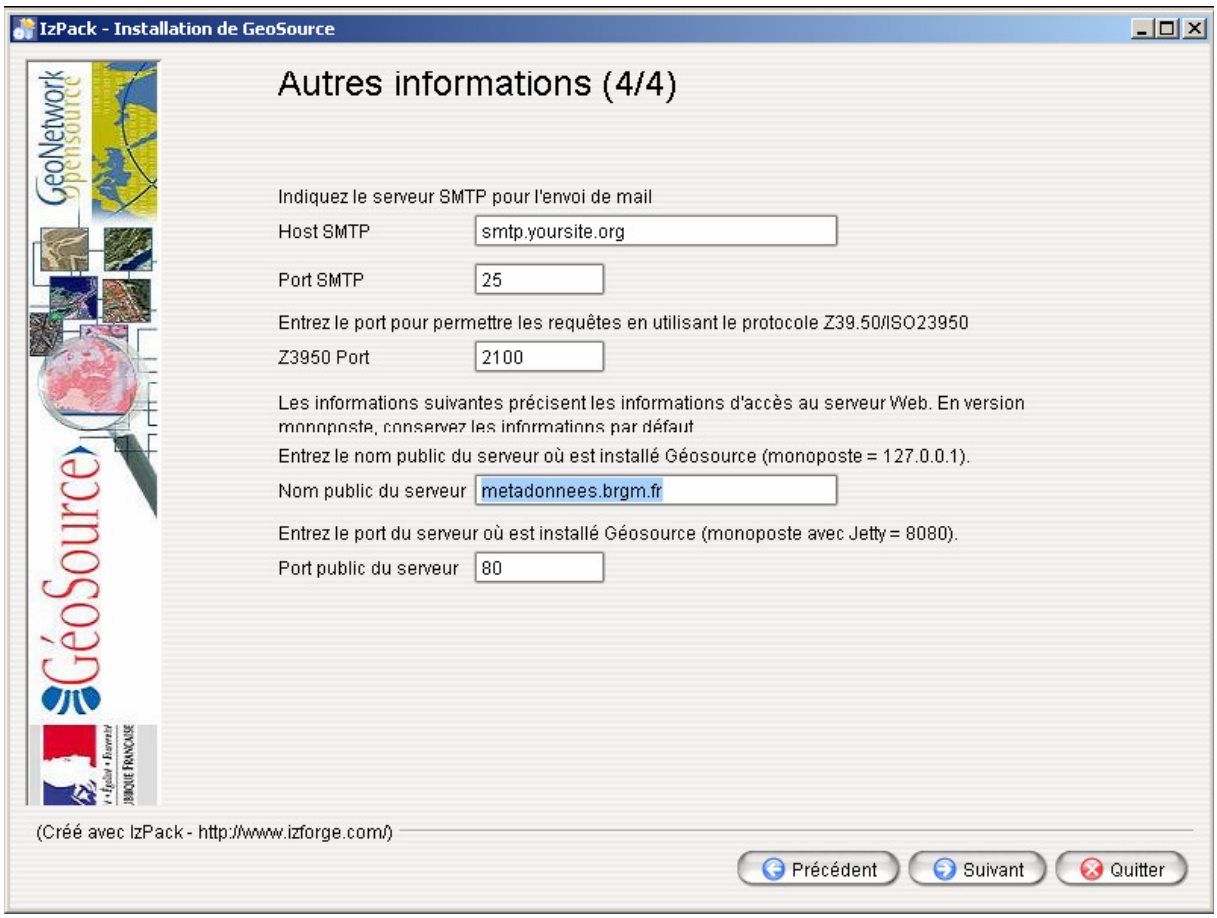
L'installation de Géosource sous Tomcat avec l'installeur est la suivante :
- Lors du choix des paquetages, désélectionnez le serveur Jetty :


IzPack - Installation de GeoSource
Choisissez ici les paquetages à installer :
Note : les paquetages grisés ne sont pas optionnels.
Core 30,08 MB
Sample metadata 45,15 KB
Integrated Jetty server 5,76 MB
Sources 950,58 KB
JavaDoc 3,54 MB
Description
A installer obligatoirement pour une utilisation en version monoposte. En version serveur, peut être remplacé par un serveur d'application spécifique
Espace total requis : 30,12 MB
Espace disque disponible : > 2 GB
(Créé avec IzPack - http://www.izforge.com/)
Précédent
Suivant
Quitter

Dans l'écran suivant, indiquez les informations sur le nom d'accès à la servlet (par défaut Geonetwork).


IzPack - Installation de GeoSource
Informations générales (1/4)
Entrez le nom de votre organisation
Nom de l'organisation
brgm
Entrez le nom d'accès à Géosource depuis votre navigateur (servlet). Il est conseillé de laisser la valeur par défaut.
Nom d'accès à Géosource (servlet name)
geosource
Entrez le port d'accès à Géosource depuis votre navigateur (si utilisation de Jetty)
Port d'accès
8080
Entrez un identifiant unique pour votre site Géosource. Eviter 'geosource' pour préférer le nom de votre organisation
Identifiant site Géosource
brgm
Entrez l'adresse de votre réseau Intranet. Géosource l'utilise pour différencier les utilisateurs externes (Internet) et les utilisateurs internes (Intranet)
Adresse réseau
127 - 0 - 0 - 0
Entrez le masque réseau.
255 - 255 - 0 - 0
(Créé avec IzPack - http://www.izforge.com/)
Précédent
Suivant
Quitter

Cette information vous sera nécessaire pour la configuration Tomcat
- Dans l'étape 4/4 , indiquez les paramètres d'accès Web à votre serveur Web, c'est-àdire l'adresse du serveur public et le port public du serveur :

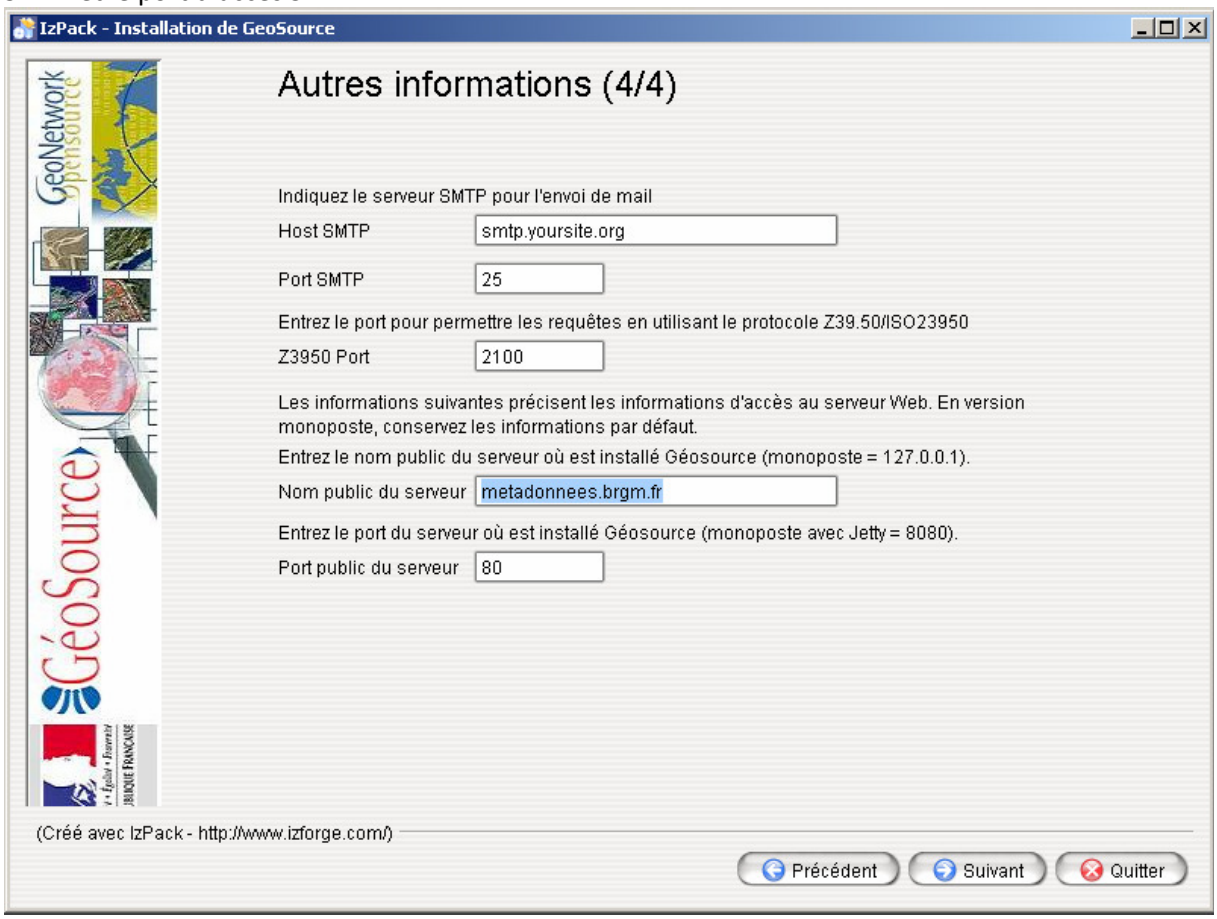

IzPack - Installation de GeoSource
Autres informations (4/4)
Indiquez le serveur SMTP pour l'envoi de mail
Host SMTP
smtp.yoursite.org
Port SMTP
25
Entrez le port pour permettre les requêtes en utilisant le protocole Z39.50/ISO23950
Z3950 Port
2100
Les informations suivantes précisent les informations d'accès au serveur Web. En version monoposte, conservez les informations par défaut
Entrez le nom public du serveur où est installé Géosource (monoposte = 127.0.0.1).
Nom public du serveur
metadonnees.brgm.fr
Entrez le port du serveur où est installé Géosource (monoposte avec Jetty = 8080).
Port public du serveur
80
(Créé avec IzPack - http://www.izforge.com/)
Précédent
Suivant
Quitter

Suite à l'installation, les opérations suivantes DOIVENT être réalisées :
1 / Vous devez au préalable faire le lien entre Géosource et Tomcat. Modifiez le fichier %TOMCAT-DIR%/conf/server.xml avec les lignes suivantes à l'intérieur de la balise <Host> :
<Context
path="/Geonetwork" < !-- Nom de la servlet -->
docBase="[Répertoire d'installation Géosource]\web"
crossContext="false"
debug="0"
reloadable="false" />
Une solution alternative est de copier le répertoire [Répertoire d'installation Géosource]/web dans le répertoire tomcat/webapps, puis de renommer ce répertoire web selon le nom de la servlet entrée dans l'étape 4 de l'installeur (généralement Geonetwork)
2 / Vous devez ensuite changer les librairies tomcat XML en copiant (et remplacant) les fichiers jar contenues dans [Répertoire d'installation Géosource]/tomcat/common/endorsed et en les copiant dans [Répertoire d'installation Tomcat]/common/endorsed
3 / Après l'installation, il est vivement recommandé d'augmenter la mémoire allouée à Tomcat.

Modifier le fichier catalina.bat avec
"set CATALINA_OPTS=-Xmx512m"
ou catalina.sh avec
"set CATALINA_OPTS=-Xmx512m"
4 / Relancer Tomcat.

5.5. LE SERVEUR DE MAIL

Géosource peut envoyer des mails depuis un formulaire de saisie par l'utilisateur. Pour cela, un serveur de mail doit être mis à disposition de l'application.
Sa configuration se réalise durant l'étape 4 de l'installeur en indiquant le nom du serveur de mail SMTP et le port d'accès SMTP.


IzPack - Installation de GeoSource
Autres informations (4/4)
Indiquez le serveur SMTP pour l'envoi de mail
Host SMTP
smtp.yoursite.org
Port SMTP
25
Entrez le port pour permettre les requêtes en utilisant le protocole Z39.50/ISO23950
Z3950 Port
2100
Les informations suivantes précisent les informations d'accès au serveur Web. En version monoposte, conservez les informations par défaut.
Entrez le nom public du serveur où est installé Géosource (monoposte = 127.0.0.1).
Nom public du serveur
metadonnees.brgm.fr
Entrez le port du serveur où est installé Géosource (monoposte avec Jetty = 8080).
Port public du serveur
80
(Créé avec IzPack - http://www.izforge.com/)
Précédent
Suivant
Quitter

Config Harvesting

MOISSONNAGE PAR DES CATALOGUES EXTERNES

Il est possible d'autoriser le moissonnage de Géosource Web par un catalogue extérieur (technique de harvesting).
Pour cela, plusieurs approches sont possibles :

Moissonnage entre serveur Géosource / Geonetwork. Dans ce cas, le système utilise directement les services d'appel de Géosource / Geonetwork.

Moissonnage entre serveur Géosource / Geonetwork

L'ouverture du service de moissonnage entre serveur Géosource / geonetwork se fait en suivant
les opérations suivantes :

1. Editer le fichier [Répertoire d'installation]\web\xml\user-profiles.xml
2. Supprimer les commentaires du service xml.metadata.get. Vous devez déclarer le service comme ceci : <allow service="xml.metadata.get"/>
3. Relancer le serveur d'application pour qu'il prenne en compte la modification
4. Vérifier que le service est actif en invoquant l'url :

http://localhost:8080/Geonetwork/srv/fr/xml.metadata.get?id=70 ou id est un identifiant

GeoNetwork - Requêtes vers un serveur distant

Lorsque vous faîtes une "Recherche distante" (Remote search), vous pouvez voir une liste de serveurs distants apparaître.

Cette liste est définié dans **[Répertoire d'installation]/web/xml/repositories.xml**

Elle contient déjà liste de serveurs que vous pouvez interroger à distance.

Chaque serveur est déclaré via 3 balises XML. Prenons l'exemple d'un serveur de cette liste :

```
<Collection collection_dn="198.102.62.179:6668/BrowseMetadata"
collection_name="Geography Network"/>
```

```
<Repository repository_dn="198.102.62.179:6668/BrowseMetadata" name="Geography Network" type="Z3950" can_multiplex_sessions="no" >

<RepositoryProperty name="ServiceHost" value="clearinghouse2.esri.com" />

<RepositoryProperty name="ServicePort" value="6668" />

<RepositoryProperty name="service_short_name" value="BrowseMetadata" />

<RepositoryProperty name="default_record_syntax" value="xml" />

<RepositoryProperty name="default_element_set_name" value="s" />

<RepositoryProperty name="full_element_set_name" value="f" />

<RepositoryProperty name="logo_src" value="http://www.k-int.com/collect_ico/sif.gif" />

</Repository>
```

<Instance instance_dn="198.102.62.179:6668/BrowseMetadata"

collection_dn="198.102.62.179:6668/BrowseMetadata"

repository_dn="198.102.62.179:6668/BrowseMetadata"

local_name="BrowseMetadata" />

Lorsque l'on souhaite ajouter un serveur distant, il suffit de modifier les éléments en couleur :

- l'adresse IP du serveur distant
- le port sur lequel écoute le Z39.50
- le nom du serveur
- le répertoire d'installation de GeoNetwork sur ce serveur
- le nom qui apparaît dans la liste des serveurs distants

Admettons que nous souhaitions ajouter un nouveau serveur à notre liste afin de pouvoir requêter dessus. Prenons par exemple un serveur avec les caractéristiques suivantes :

- adresse IP : **10.25.184.56**
- port z3950 : **2100**
- nom du serveur **: vmgeo.ird.fr**
- répertoire d'install : **geonetwork**
- nom dans la liste des serveurs distants : **notre_serveur**

Il faudra donc ajouter les lignes suivantes dans le fichier repositories.xml :

<Collection collection_dn="10.25.184.56:2100/geonetwork"
collection_name="notre_serveur"/>

```
<Repository repository_dn="10.25.184.56:2100/geonetwork" name="notre_serveur"
type="Z3950" can_multiplex_sessions="no" >

<RepositoryProperty name="ServiceHost" value="vmgeo.ird.fr" />

<RepositoryProperty name="ServicePort" value="2100" />

<RepositoryProperty name="service_short_name" value="geonetwork" />

<RepositoryProperty name="default_record_syntax" value="xml" />

<RepositoryProperty name="default_element_set_name" value="s" />

<RepositoryProperty name="full_element_set_name" value="f" />

<RepositoryProperty name="logo_src" value="http://www.k-int.com/collect_ico/sif.gif" />

</Repository>


<Instance instance_dn="10.25.184.56:2100/geonetwork"
collection_dn="10.25.184.56:2100/geonetwork"

repository_dn="10.25.184.56:2100/geonetwork"

local_name="geonetwork" />
```

Il suffit ensuite de redémarrer le serveur Jetty ou Tomcat

ATTENTION AUX PROBLEMES

NB : il faut bien évidemment que, sur le serveur distant, le port du Z39,50 soit ouvert et atteignable (qu'il n'y ait pas un firewall qui en bloque l'accès).